

#5

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/824,735

DATE: 09/07/2001

TIME: 14:11:35

Input Set : A:\205645US20.txt

Output Set: N:\CRF3\09072001\I824735.raw

ENTERED

3 <110> APPLICANT: ZHU, JIAN-KANG  
4 LIU, JIPING  
5 ISHITANI, MANABU  
6 HALFTER, URSULA  
7 KIM, CHEOL-SOO  
9 <120> TITLE OF INVENTION: PROTEINS AND DNA RELATED TO SALT TOLERANCE IN PLANTS  
11 <130> FILE REFERENCE: 205645US20  
13 <140> CURRENT APPLICATION NUMBER: 09/824,735  
C--> 14 <141> CURRENT FILING DATE: 2001-08-17  
16 <150> PRIOR APPLICATION NUMBER: US 60/824,735  
17 <151> PRIOR FILING DATE: 2000-04-04  
19 <160> NUMBER OF SEQ ID NOS: 16  
21 <170> SOFTWARE: PatentIn version 3.1  
23 <210> SEQ ID NO: 1  
24 <211> LENGTH: 5144  
25 <212> TYPE: DNA  
26 <213> ORGANISM: Arabidopsis thaliana  
28 <220> FEATURE:  
29 <221> NAME/KEY: CDS  
30 <222> LOCATION: (2004)..(2168)  
31 <223> OTHER INFORMATION:  
34 <220> FEATURE:  
35 <221> NAME/KEY: CDS  
36 <222> LOCATION: (2255)..(2317)  
37 <223> OTHER INFORMATION:  
40 <220> FEATURE:  
41 <221> NAME/KEY: CDS  
42 <222> LOCATION: (2416)..(2487)  
43 <223> OTHER INFORMATION:  
46 <220> FEATURE:  
47 <221> NAME/KEY: CDS  
48 <222> LOCATION: (2573)..(2680)  
49 <223> OTHER INFORMATION:  
52 <220> FEATURE:  
53 <221> NAME/KEY: CDS  
54 <222> LOCATION: (2778)..(2852)  
55 <223> OTHER INFORMATION:  
58 <220> FEATURE:  
59 <221> NAME/KEY: CDS  
60 <222> LOCATION: (2921)..(3100)  
61 <223> OTHER INFORMATION:  
64 <220> FEATURE:  
65 <221> NAME/KEY: CDS  
66 <222> LOCATION: (3225)..(3314)  
67 <223> OTHER INFORMATION:  
70 <220> FEATURE:  
71 <221> NAME/KEY: CDS

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/824,735

TIME: 14:11:35

Input Set : A:\205645US20.txt

Output Set: N:\CRF3\09072001\I824735.raw

```

72 <222> LOCATION: (3418)..(3540)
73 <223> OTHER INFORMATION:
76 <220> FEATURE:
77 <221> NAME/KEY: CDS
78 <222> LOCATION: (4052)..(4171)
79 <223> OTHER INFORMATION:
82 <220> FEATURE:
83 <221> NAME/KEY: CDS
84 <222> LOCATION: (4400)..(4516)
85 <223> OTHER INFORMATION:
88 <220> FEATURE:
89 <221> NAME/KEY: CDS
90 <222> LOCATION: (4735)..(4791)
91 <223> OTHER INFORMATION:
94 <220> FEATURE:
95 <221> NAME/KEY: CDS
96 <222> LOCATION: (4879)..(4953)
97 <223> OTHER INFORMATION:
100 <220> FEATURE:
101 <221> NAME/KEY: CDS
102 <222> LOCATION: (5038)..(5130)
103 <223> OTHER INFORMATION:
106 <400> SEQUENCE: 1
107 tttttggcgg gaaaatctcg ggtttacgtt tttggcgga aaatctcgtg tttacgtttt      60
109 tggcgggaaa atctcgggtt tacgtttttt gcgagaaaat cacgggttta cttttttttg      120
111 cgggaaaatc acggatttat gttttttggt ggaaaaatta cgagtttact ttttctcaat      180
113 ttcattgctt gtatttttaa gaaatttgga aaaatattaa ttttattaaa ttggtttaga      240
115 tgtgttggtt aaacctaaat tggcattggt ttagagattt tagttggttt tattcaattt      300
117 tacaaaattt gatgggttaa ttggataaac catggaaacc attaaccatt acaacctaac      360
119 tcattttact catcaaacca attgactcat caactcattt gactcatcaa ctcatgtgag      420
121 tcaaaaattt caactcatta gggttcatgg gttgagttga gttgagttga cccatgaatt      480
123 ttgaccttatt ttgacacccc tacatatgat cataagttaa taatcaaaaa ttactattga      540
125 taacttttta acggaattgt tttgtaagtt tcattttacgt tatttatata aaaaaacata      600
127 atgcaaaagt actaatgtat agttattttt atttttaata atgcaaaaat attactgtaa      660
129 tacttttttca ttcttatcaa ttttttttct tatattcaat tagccaccaa cacctacatt      720
131 tcatattttct cttcacattc attatttttct tgctcttgta ttcccttctta tcatcttcat      780
133 tgccaattttt tcattgacat tgtcatcggt actttgtatg tatgattttt gaacatttaa      840
135 tgtgaatccc aaccgactat aggactatca agaagttttc aaacttttta aaaaagatct      900
137 tgaaccttta aagcaaatcc caaccgacta tagaaccata gtcctcttcc ttgttgatga      960
139 agctcttctc gtgccggcga aaatctaggc cataaaaagcc tcttcaacat cacctagtat      1020
141 attgaccgtg accatctttt tgaccattgc tttgtgaatg aaccgtcgat aaaccgtgtt      1080
143 atcacttacg ccaaattttt ccctagtgtt tgtttccaaa ctctcacgaa tccttatcga      1140
145 actttttttat atatcacttt gtagcattgg aaagtatctt tgtatgcttt gtcttaaaact      1200
147 tagacatcct tgttctcttg gtttttttga accttgcttg acttaaatga agttaaaaaat      1260
149 ttgtagttaa aaatagaaaa ttttactaat ttgagttcga ttaatcatag tctagataat      1320
151 ttgaaaaaaa ttaaaataaa ttttgaaaat actatatgtt ttttttttaa aaataattta      1380
153 ctaaattgat aagtgatatt agattagttt ttttttctt ttttaaattt tgaaaacctc      1440
155 acaattatta aattgaaaga ttcaaaatgc aatgttagtt tttaaaagtt taatcaccca      1500
157 aacggataat tgacctgaac gactaattca ggtcgtatac gggtacagtc aaataacccg      1560

```

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/824,735

TIME: 14:11:35

Input Set : A:\205645US20.txt

Output Set: N:\CRF3\09072001\I824735.raw

```

159 acccgaaatg ctcaacggat ctgcacggac gttcgcacg acgattcacg gctttcgcga 1620
161 atcgcatcac gagccttcct tctctcctac gcctctttca tcaacccttc tctgcgaatc 1680
163 caattctggt atttcacgat tctctttcga ctacgcccaa tcgcaattcc agccgtacga 1740
165 attttatata ttgattatga tcttgatcct taccttttcg tcgtttcctt gatcgaattt 1800
167 atgttgattt atgctgcgaa atcaaaatta tggatacaga gtaattttgt atatatggat 1860
169 gtgtaggagc taaaattagc aaacaatacg taaagtaatt gaaatcgaaa tcataaacgt 1920
171 ttaaggaaaag aggtttttac taagtctctg aataatctga ttgatagctt gtggtcaatg 1980
173 gatcagataa aagtttgtaa aga atg aca aag aaa atg aga aga gtg ggc aag 2033
174                                     Met Thr Lys Lys Met Arg Arg Val Gly Lys
175                                     1                               5                               10
177 tac gag gtt ggt cgc aca ata ggt gaa gga acc ttt gct aag gtt aag 2081
178 Tyr Glu Val Gly Arg Thr Ile Gly Glu Gly Thr Phe Ala Lys Val Lys
179                                     15                               20                               25
181 ttt gcg agg aac aca gac act ggt gat aat gta gcc atc aaa att atg 2129
182 Phe Ala Arg Asn Thr Asp Thr Gly Asp Asn Val Ala Ile Lys Ile Met
183                                     30                               35                               40
185 gct aag agt aca ata ctt aag aac aga atg gtt gat cag gtatgttctg 2178
186 Ala Lys Ser Thr Ile Leu Lys Asn Arg Met Val Asp Gln
187                                     45                               50                               55
189 gattgttttt tacatggaaa ctaagggtgt tgcgtcaatg gtatgatctt tgatttcggt 2238
191 taaagctctt ttacag ata aaa aga gag ata tct ata atg aag att gtt cgt 2290
192                                     Ile Lys Arg Glu Ile Ser Ile Met Lys Ile Val Arg
193                                     60                               65
195 cac ccg aac ata gtg agg ttg tat gag gtatgtttgt ttgtttccat 2337
196 His Pro Asn Ile Val Arg Leu Tyr Glu
197                                     70                               75
199 gcactctgcga aattttatct ctgaagtgtt tttgcatcat tgttcttctg ttgttttttt 2397
201 gtgattttcc cgatgtag gtg ttg gcg agt cct tcg aaa ata tat ata gtt 2448
202                                     Val Leu Ala Ser Pro Ser Lys Ile Tyr Ile Val
203                                     80                               85
205 ttg gag ttt gtg aca gga gga gag ctc ttt gat aga att gtacggaact 2497
206 Leu Glu Phe Val Thr Gly Gly Glu Leu Phe Asp Arg Ile
207                                     90                               95                               100
209 tccatacttg taggcagcgt ccattagtta aaacctctct acttaatttt ttaatatatg 2557
211 aaatctttca tgcag gtt cat aaa ggg agg ctt gaa gaa agt gag tct cgg 2608
212                                     Val His Lys Gly Arg Leu Glu Glu Ser Glu Ser Arg
213                                     105                               110
215 aaa tac ttt caa cag ctt gta gat gct gtt gct cat tgt cac tgc aag 2656
216 Lys Tyr Phe Gln Gln Leu Val Asp Ala Val Ala His Cys His Cys Lys
217                                     115                               120                               125
219 ggt gtt tac cac cgt gac cta aag gtaaagacgt gtttttgttt accaatattc 2710
220 Gly Val Tyr His Arg Asp Leu Lys
221                                     130                               135
223 ctcaagaatat ctcaactgcgt tgcaatccag acttgatatt tttgtgtcgc tatgttatgt 2770
225 tatctag cca gaa aat ctt tta ctc gat aca aat gga aat ctg aag gtt 2819
226                                     Pro Glu Asn Leu Leu Leu Asp Thr Asn Gly Asn Leu Lys Val
227                                     140                               145                               150
229 tcg gat ttc gga ctc agt gca ttg cct cag gaa gtaagtgtc ttatctctgc 2872
230 Ser Asp Phe Gly Leu Ser Ala Leu Pro Gln Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/824,735

DATE: 09/07/2001

TIME: 14:11:35

Input Set : A:\205645US20.txt

Output Set: N:\CRF3\09072001\I824735.raw

```

231          155          160
233 ttcagcagtc tgcttacgtg gtcattaact tggtatatac tcaatcag gga gta gaa 2929
234                                     Gly Val Glu
237 ctt ctg cgt acc aca tgt gga act ccg aac tat gta gct cca gag gta 2977
238 Leu Leu Arg Thr Thr Cys Gly Thr Pro Asn Tyr Val Ala Pro Glu Val
239 165          170          175          180
241 ctt agt gga cag ggt tac gat ggt tca gca gct gat att tgg tct tgc 3025
242 Leu Ser Gly Gln Gly Tyr Asp Gly Ser Ala Ala Asp Ile Trp Ser Cys
243          185          190          195
245 ggg gtt att ctt ttc gtt ata ttg gct gga tat tta cct ttt tcc gag 3073
246 Gly Val Ile Leu Phe Val Ile Leu Ala Gly Tyr Leu Pro Phe Ser Glu
247          200          205          210
249 acg gat ctt cca ggg ttg tac aga aaa gtaagtaaca tatctttcgg 3120
250 Thr Asp Leu Pro Gly Leu Tyr Arg Lys
251          215          220
253 gaagaaatca tgaattcctt gtcattgctt ttgtcaaacc gtttattgat ttgggtttgc 3180
255 aatttcaccc ttagttttga gcttttacac attgttattt acag ata aat gca gca 3236
256                                     Ile Asn Ala Ala
257          225
259 gag ttt tct tgt cca ccg tgg ttt tcc gca gaa gtg aag ttt tta ata 3284
260 Glu Phe Ser Cys Pro Pro Trp Phe Ser Ala Glu Val Lys Phe Leu Ile
261          230          235          240
263 cat agg ata ctt gac ccc aat ccc aaa aca gtgagtattt tgctttgttc 3334
264 His Arg Ile Leu Asp Pro Asn Pro Lys Thr
265          245          250
267 tctcctagct atcagggtttt ggtgatattt aatgttctag taattatatc tgtttatcta 3394
269 ttattgtttc tcaattagag cag cgt att caa att caa gga atc aag aaa gat 3447
270                                     Arg Ile Gln Ile Gln Gly Ile Lys Lys Asp
271          255          260
273 cct tgg ttc aga tta aat tat gtg cct ata cga gca agg gaa gaa gaa 3495
274 Pro Trp Phe Arg Leu Asn Tyr Val Pro Ile Arg Ala Arg Glu Glu Glu
275          265          270          275
277 gaa gtg aat ttg gat gat att cgt gca gtt ttt gat gga att gag 3540
278 Glu Val Asn Leu Asp Asp Ile Arg Ala Val Phe Asp Gly Ile Glu
279          280          285          290
281 gtttgtgctt cgccttcatt atatgctctt tgctggtcaa ttccatttaa atgttaagat 3600
283 ctcttaggaa cgtttgatg accaagaaga agtgtttgct acaggataga acaaaatagt 3660
285 aggcatgtgt tagttaccaa acctgtaaac tgcttcttta ttcaattcgc caaaccatag 3720
287 accttaggaa gacttagatt tacaagagaa ttctctattc tcgacaaaa accctagaca 3780
289 aaatccagaa tacccttagg gctaattaca atgttcatgt acctatcaat atatatctcc 3840
291 tgctaagatt gtcttagctt tggtagacc tagatatata tagataccgt aatttctaata 3900
293 gcatatttag aatgttttac ttaaaactcag gctcttgctc ttctaaaact tgtacttcaa 3960
295 ttgttaaact aaaacctcag tatctgtctt agctaaagtt acttttactt gtttttcatt 4020
297 aagttgacct gtcaattgca cttgttcaca g ggc agt tat gta gcg gag aat 4072
298                                     Gly Ser Tyr Val Ala Glu Asn
299          295
301 gta gag aga aat gat gaa ggg ccc ctg atg atg aat gcc ttt gag atg 4120
302 Val Glu Arg Asn Asp Glu Gly Pro Leu Met Met Asn Ala Phe Glu Met
303 300          305          310          315

```

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/824,735

TIME: 14:11:35

Input Set : A:\205645US20.txt

Output Set: N:\CRF3\09072001\I824735.raw

```

305 att acc tta tca caa ggc tta aat tta tct gca cta ttt gac agg cga      4168
306 Ile Thr Leu Ser Gln Gly Leu Asn Leu Ser Ala Leu Phe Asp Arg Arg
307                               320                               325                               330
309 cag gtagtacctg attttctatt actggtcata gagatctcca tttcgaataa      4221
310 Gln
313 aagaatgtcg gtagcatcta ttcttcagac tgcccgtttt gactgcctta tgatgctgtg      4281
315 ttcttagttt gttataataa ctataagttc attagatgat tggttgcatg gcattagtag      4341
317 atacaaatgg aatccaaaat gttcctgcat attgatggct gatcctttga tctcgag      4399
319 gat ttt gtt aaa agg caa acc cgt ttt gtt tct cga agg gaa cct agt      4447
320 Asp Phe Val Lys Arg Gln Thr Arg Phe Val Ser Arg Arg Glu Pro Ser
321                               335                               340                               345
323 gag ata att gct aac att gag gct gta gcg aac tca atg ggt ttt aag      4495
324 Glu Ile Ile Ala Asn Ile Glu Ala Val Ala Asn Ser Met Gly Phe Lys
325                               350                               355                               360
327 tct cat aca cga aac ttc aag gtaacgaatt cctagcatat tacacttattc      4546
328 Ser His Thr Arg Asn Phe Lys
329 365                               370
331 acagagatta tgcattattht taaaactctc aactgttaaa cgcatgtgta gatagattga      4606
333 taagattgac aaggaaactt agtttatatc tctggcgttc aaaaacgaaa gtcctagtg      4666
335 gaattatcat tttaatggt agcagagagt acaattgtta tgatttgta cgtctatgtg      4726
337 ctcaacag aca agg ctc gag gga tta tct tcg atc aag gcc gga cag tta      4776
338 Thr Arg Leu Glu Gly Leu Ser Ser Ile Lys Ala Gly Gln Leu
339                               375                               380                               385
341 gct gtt gtg ata gag gtaattattg cttgttgga ttgtaataata agtttgcttt      4831
342 Ala Val Val Ile Glu
343                               390
345 gcttcagttt aagggtatct agcaaattga aattaaccta catgcag att tac gag      4887
346                               Ile Tyr Glu
349 gtg gca cca tcg ctt ttc atg gta gac gta aga aag gct gct ggt gaa      4935
350 Val Ala Pro Ser Leu Phe Met Val Asp Val Arg Lys Ala Ala Gly Glu
351                               395                               400                               405
353 act ctt gaa tat cac aag gtttataaat atatatccaa taacaatagt      4983
354 Thr Leu Glu Tyr His Lys
355 410                               415
357 tgcatcatta ctgtgttgcg gattagagtg atattttgtt ttgtgggtatc gcag ttc      5040
358                               Phe
361 tac aag aag cta tgt tcg aaa ctg gaa aac ata ata tgg agg gca aca      5088
362 Tyr Lys Lys Leu Cys Ser Lys Leu Glu Asn Ile Ile Trp Arg Ala Thr
363                               420                               425                               430
365 gaa gga ata cca aag tca gag att ctc aga aca atc acg ttt      5130
366 Glu Gly Ile Pro Lys Ser Glu Ile Leu Arg Thr Ile Thr Phe
367                               435                               440                               445
369 tgatcccaac ttaa      5144
372 <210> SEQ ID NO: 2
373 <211> LENGTH: 446
374 <212> TYPE: PRT
375 <213> ORGANISM: Arabidopsis thaliana
377 <400> SEQUENCE: 2
379 Met Thr Lys Lys Met Arg Arg Val Gly Lys Tyr Glu Val Gly Arg Thr

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/824,735

DATE: 09/07/2001

TIME: 14:11:36

Input Set : A:\205645US20.txt

Output Set: N:\CRF3\09072001\I824735.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date